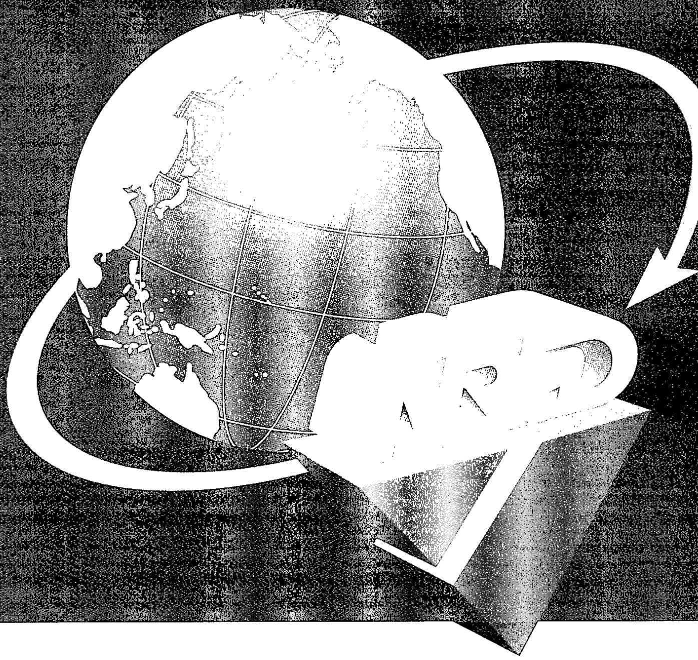


Technical Report 1728
June 1996

Evaluating Collaborative Technologies for Command and Control Teams

J. W. Gwynne B. A. Feher R. W. Obermayer R. J. Smillie J. M. Linville N. J. Heacox

Naval Command, Control and
Ocean Surveillance Center
RDT&E Division

San Diego, CA
92152–5001

Technical Report 1728
June 1996

Evaluating Collaborative Technologies for Command and Control Teams

J. W. Gwynne
B. A. Feher
R. W. Obermayer
R. J. Smillie
J. M. Linville
N. J. Heacox

NAVAL COMMAND, CONTROL AND OCEAN SURVEILLANCE CENTER
RDT&E DIVISION
San Diego, California 92152-5001

H. A. WILLIAMS, CAPT, USN
Commanding Officer

R. C. KOLB
Acting Executive Director

ADMINISTRATIVE INFORMATION

The work reported in this report was performed for the Collaborative Technologies Project Team (Code 44210) of the Simulation and Human Systems Technology Division (Code 44) of the Command and Control Department (Code 40) of the Research, Development, Test and Evaluation Division of the Naval Command, Control and Ocean Surveillance Center by Pacific Science and Engineering Group, Inc. as Delivery Order 36 under Contract N66001–90–D–0127. Funding was provided by the Office of Naval Research (ONR 342, Dr. Stanley Collyer) under program element 0602233N RM33H20 (Human Factors Technology). The Collaborative Technologies Project Team was led by Robert Smillie. This report covers work from October 1994 to January 1996.

Released by
R. J. Smillie, Head
Collaborative Technologies Branch

Under authority of
J. D. Grossman, Head
Simulation and Human Systems
Technology Division

ACKNOWLEDGMENTS

The authors acknowledge the contributions of Robert Smillie and Matthew Scallon of NRaD, as well as Richard W. Obermayer and J. Martin Linvolle of Pacific Science and Engineering Group, Inc. in the setup of the laboratory and development of the scenarios and measurement instruments. In addition, we wish to express appreciation to the members of the following teams and their leaders who participated in the simulations:

- ONR Naval Reserve Unit
- Australian Navy Team
- CRUDESGRU 5
- COMCRUDESGRU 5 — CDR John Walters
- COMCRUDESGRU 5 — ENS Mike Rutledge
- COMCRUDESGRU 5 — LT Paul McKeon
- C2 Warfare Center — LT Alan Philpott
- COMCARGRU 7 — CDR Tom Grant
- TACTRAGRUPAC — LT Privateer

SB

SUMMARY

This document describes the U. S. Navy's initial demonstration of the COllaborative LABoratory (COLAB) during a distributed situation assessment task. Five command and control warfare (C^2W) teams, each consisting of three naval officers and a commander, participated in an exercise scenario involving four vignettes that required the team to generate a situation assesment report (SITREP) based on shared information. The evaluation team compared the quality, completeness, and consistency of team performance using collaborative technology with more traditional voice and text communications methods. No advantage could be attributed solely to the collaborative technology. Differences in the frequency and pattern of communications across the traditional and collaborative methods created nonmeaningful influences on the major outcome measures. However, these results led to improved understanding of team collaborative behavior and the creation of a systematic approach for assessing the role of collaborative technology on team performance. As part of this approach, the evaluation team developed a set of analytic techniques for evaluating data collected in collaborative studies. These techniques can be applied in future experiments using the COLAB to evaluate collaborative processes and technologies in Navy C^2W operations.

CONTENTS

INTRODUCTION

Collaborative software technologies (groupware) facilitate the communication and exchange of information among team members who are working together to complete a shared task. These individuals may either be located in the same general vicinity or dispersed over a wide geographic region. Before the full potential of collaborative software technologies can be realized, however, research is needed to determine how these technologies can be optimally applied to specific task environments.

In order to explore these issues in terms of navy command and control warfare (C^2W) tasks, the Naval Command, Control and Ocean Surveillance Center RDT&E Divison (NRaD) established a COllaboration LABoratory (COLAB). The mission of the COLAB is to support the systematic empirical study of information processing by teams using advanced computer networking and software technologies in domains of interest to the U.S. Navy. The COLAB is used to assess collaborative technologies for C^2W teams during simulated planning and selected operational activities. Specific tasks in this laboratory enable collaborative technology demonstrations, evaluations, workflow analyses, and assessments of distributed team activities such as collaborative problem solving, situation assessment, and planning using collaborative technologies.

This document evaluates performance of five four-person teams of naval officers on a simulated noncombatant evacuation operation (NEO) with attributes identical to operational C^2W situation assessment tasks performed by Navy battle groups and joint task forces. Performance of the teams was evaluated primarily in terms of three aspects of their Situation Assessment Reports (SITREPs):

1. Completeness of their SITREPs following the vignette;
2. Quality of their SITREPs when compared to independent expert judgment; and
3. Consistency of SITREPs among team members.

Additionally, all team communications (including verbalizations, e-mail, and Whiteboard) were analyzed for their relationship to scenario events. A timeline analysis of team performance was conducted using a sequential data analysis technique, and post-experiment questionnaires were evaluated to determine subjects' experience using computers and the utility and usability of the software tools.

PURPOSE

The purpose of this study was to demonstrate NRaD's COLAB. This laboratory was used to evaluate design, development, test, evaluation, and application of collaborative technologies to navy operational tasks as performed by experienced naval personnel. As part of this effort, an initial study investigated factors that contribute to the effectiveness of collaborative software tools, with a particular focus on interface design variables.

OBJECTIVES

Specific objectives of this study were to:

1. Demonstrate the feasibility of examining collaborative technology issues in a simulated navy operational setting using commercial, off-the-shelf groupware;

2. Develop data analysis techniques that permit evaluation of team member collaboration with emphasis on software tool interface features and realistic operational objectives;

3. Determine the impact of collaborative tools on the outcome of a team situation assessment task;

4. Identify the interface features of collaborative tools that influence their effectiveness;and

5. Identify team-level tasks that can benefit most from collaborative technologies.

DESCRIPTION OF COLAB FACILITY: DATA COLLECTION AND ANALYSIS

LABORATORY DESIGN

The COLAB, developed and installed at NRaD, enables assessment of the type and extent of collaboration possible when using various combinations of collaborative software tools. The COLAB allows controlled manipulation of variables during C^2W team collaboration and precise response measurement for both individual team members and an overall team. Figure 1 shows the physical layout of the COLAB, which consists of a control room and a partitioned work room with four workstations. The control and work rooms are connected by digital and voice communication networks.

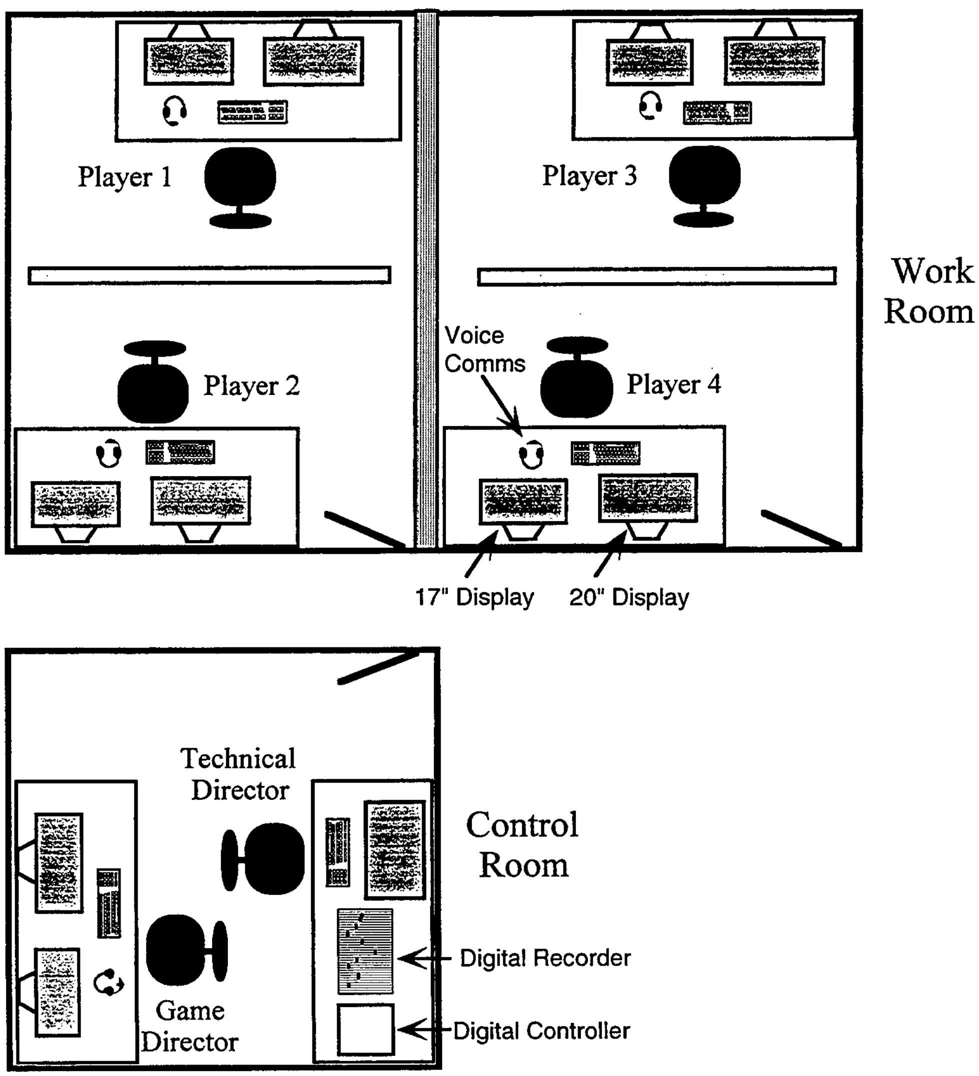


Figure 1. Floor plan of COLAB experimental facility.

The control room contains two workstations for controlling the experiment and capturing data for analysis and after-action reviews. The Game Director directs experiment start/pause/stop and provides geopolitical summaries. During the experiment, the Game Director acts as all other participants, e.g., bridge, Tactical Action Officer (TAO), Battle Group Commander (GB), and Aircraft Warning and Control System (AWACS). The Technical Director operates the data collection devices and scenario driver, monitors system operation, and selects and initiates scenarios. Voice recording equipment, communications hardware, and software are used to generate and monitor data produced at each team workstation during simulated tactical scenarios. Within the work room, partitions separate player workstations during experimental sessions to prevent information exchange among team members.

Each workstation consists of a Macintosh 660AV computer with two video monitors, a hard disk, a keyboard, a mouse, and enough RAM to run up to six applications simultaneously. All six workstations are connected by an Ethernet network that transfers control commands, scenario information, data, and eleectronic messages. An independent audio network provides voice communications between stations. This audio network can be disabled, enabled in broadcast mode (everyone hears everyone else), or enabled in direct mode (yielding only point-to-point communications).

The experiment control software includes data logging of actions by team members, logging of human-computer interaction events across all system applications, and recording of all communications (with time stamping to permit correlation with other team member responses). A simulation driver (SIMDRIVER) program presents a tactical situation display with alert messages at each of the four workstations. The two displays at each workstation present the tactical situation and the software tools. The tactical display includes a window that presents all stored information on a hooked track. Another window presents incoming alert messages that team members were instructed to highlight and mark as soon as they read them.

The control station is used to monitor the scenario, update satellite maps, and collect data from player workstations between vignettes. All data sent between workstations, either through the Ethernet network or by voice communications, are stored for later analysis. In addition, each keystroke or other interaction between a team member and his workstation is recorded to document the team member's activities.

The SIMDRIVER program in the control station provides tactical data to each workstation. Each workstation runs a uniquely configured program that accepts tactical data, selects data needed to perform its functions, and displays it to the operator. Each workstation displays a changing tactical map on one of its monitors. This map includes terrain, continuously updated track symbols, and land or surface installations. Tracks can be hooked to access additional information regarding their platform or affiliation, or to declutter dense clusters of symbols. Each workstation also displays workstation-specific information on the second monitor. A SITREP form is displayed continuously on each player's display. This form contains items that assess a player's understanding of the tactical situation as it develops and solicits his expectations of anticipated events (see Appendix A).

Other programs provide communications and collaborative features. An electronic mail system allows operators to compose, address, transmit, and forward text messages and attachments. A screen-capture program allows operators to capture pieces of the tactical map to be transmitted to others. A whiteboard program periodically updates a background satellite map of the mission area and allows operators to add text or to draw on the map. This whiteboard and map can, under experimental control, either be shared with all other workstations or used solely as a local resource.

COMMUNICATIONS

The COLAB uses two commercial off-the-shelf software applications with or without verbal communications to create different collaboration conditions. Team members use an electronic mail program to send text and/or graphics to each other. A whiteboard application permits operators to collaborate simultaneously through the use of color-coded text and graphics and user-specific pointers in a common work area or whiteboard. The whiteboard can feature some aspect of a task that lends itself to collaboration, such as a geographical image of the tactical situation that is updated periodically during the scenario. Team members communicate through a headphone-microphone set using either a broadcast or point-to-point mode, as determined by the evaluation design.

ANALYTIC TECHNIQUES TO ASSESS SUBJECT PERFORMANCE

Several COLAB-developed analytic techniques measure important aspects of player performance. These techniques are used to evaluate the complex datasets that are collected during the COLAB scenarios. Each technique is generic and could be applied to datasets generated by future studies that use the COLAB. Because these techniques represent novel approaches to analyzing collaborative aspects of team performance, they are discussed separately in this section.

Situation Assessment Reports (SITREPs)

Three team performance measures were derived from each team member's completed SITREP: Quality, Completeness, and Consistency. These measures capture important aspects of team performance that pertain to the nature and extent of collaboration among team members.

Quality. The Quality score provides an overall assessment of how well each team member identified important aspects of each vignette. This measure is based upon the judgment of the project subject matter expert (SME) regarding what contacts pose the greatest near-term threat. A high Quality score can be obtained by specifically mentioning each threatening contact in the SITREP.

Table 1 illustrates how a Quality score is computed based on the scoring of the SITREPs. The example in table 1 consists of six threat areas with an indication of the number of team members who mentioned each area in the SITREP. An individual's Quality score is computed by dividing the number of near-term threats identified in each SITREP by the total number of threats identified by the SME to obtain the proportion of near-term threats. Quality scores for SITREPs completed by team members in the same experimental conditions are averaged to obtain an overall Quality score for each experimental condition.

To assess team collaboration, the score of the Command and Control Warfare Commander (C^2WC) is compared to the average score for the rest of the team. The rationale for this approach is based on the fact that team members are required to provide information regarding immediate threats to the C^2WC so the C^2WC can integrate this information into an official SITREP. If team members provide the C^2WC with this information, the C^2WC should be better able to identify immediate threats and, therefore, obtain a higher Quality score than the rest of the team.

Table 1. Example calculations of Quality, Completeness, and Consistency based on the SITREP.

Threat Area	Threats	C^2WC	Crypto	Intel	EW	Agreement
1	RBF can launch SAMs from SW peninsula	X	X	X	X	3
2	PRC can intercept landing force with ASMs, SAMs, fighters; ability to control forces	X	X	X	X	3
3	PRC can intercept landing forces with PGMs	X	X	X	X	3
4	NK frigate class ships could interdict TF as it departs AOA.		X	X		1
5	NK air defense, armor, infantry, artillery could interfere with land-based activities in vicinity of PRC-NK border; SAM/SSM threat to TF				X	0
6	Russian AGI can intercept comms, etc. and pass targeting info to forces in area; intentions unknown	X		X	X	2
	Individual team member sums	4	4	5	5	
		Total Agreement				12

Quality Calculation

Threat Areas 1 and 6 were defined as near-term threats by the SME.

All team members cited Threat Area 1; all team members except Crypto cited Threat Area 6.

Quality for each team member:

C^2WC	Crypto	Intel	EW
2/2	1/2	2/2	2/2

Quality score for C^2WC = 2/2 = 1

Quality score for rest of team = (1/2 + 2/2 + 2/2)/3 = 0.83

Completeness Calculation

Total of 6 threat areas

Completeness for each team member:

C^2WC	Crypto	Intel	EW
4/6	4/6	5/6	5/6

Completeness score for C^2WC = 4/6 = 0.67

Completeness score for rest of team = (4/6 + 5/6 + 5/6)/3 = 0.78

Consistency Calculation

Numerator: Total agreement = 12

Denominator: 6 threat areas x 3 maximum agreement = 18

Consistency score = 12/18 = 0.75

Completeness. The Completeness score measures the extent to which team members identified all potential threats. A high Completeness score means that a team member identifed most (or all) of the threats identified by the SMEs.

An individual's Completeness score is computed by counting the number of different threat areas identified by that individual. (Threat areas are geographic regions containing one or more contacts that are actually or potentially hostile to one's own forces.) This value is then divided by the total number of threat areas identified by the SMEs to obtain the Completeness score. As with the Quality

score, two Completeness scores are computed, one for the C^2WC and one for the remainder of the team.

Consistency. The Consistency score measures the extent to which team members identified the same threats. In contrast to the Quality and Completeness scores, the Consistency score is a team measure and is not computed separately for the C^2WC and the rest of the team. A high Consistency score does not necessarily imply that the most relevant threats were identified; it means only that the team as a whole remained focused on a certain set of threats, whether or not they are relevant and complete.

The Consistency score is computed by counting the number of different threat areas identified by an entire team. Next, the number of degrees of freedom for each threat area is computed as one less than the number of team members who mentioned a given threat area. These values are then summed to obtain the numerator for the Consistency score. The highest agreement within an area is then multiplied by the total number of threat areas to obtain the denominator.

Timeline Analysis

MacSHAPA[1] was used to conduct timeline and frequency analysis, as well as transition analysis, which allowed empirical mapping of communication interaction patterns. MacSHAPA also was used to compare interaction sequences between team members.

Timeline analysis is a powerful approach to evaluating sequential data that were collected from several communications sources, e.g., voice, e-mail, and whiteboard. Timeline analysis permits team member responses to be temporally related to specific scenario events as well as to other team member responses. The analysis produces a detailed description of scenario stimuli and team member responses to these stimuli. Specific hypotheses can be tested by examining selected aspects of an event sequence, and combinations of variables can be investigated for their relationship to task performance measures, e.g., marking alert messages.

Results of a timeline analysis can be plotted on an individual or team basis to graphically display interaction sequences, including communication density and patterns. Timeline plots allow the information flow to be tracked throughout the team. For example, because each communication can be tagged by its related threat area, it is possible to follow the information exchanges by threat area within a vignette, and to determine which team members shared information relative to a particular threat area.

Categorization of Communications

The evaluation team developed a category rating scheme to evaluate the type and extent of information sharing among team members. The rating scheme consists of three categories: Message Type, Message Subject, and Message Content. When taken together, these categories describe the characteristics of the team's communications. Table 2 shows the hierarchical relationship among the rating scheme's first three categories. The first level, Message Type, consists of three alternatives: Request, Provide, and Reply. The second level, Message Subject, also has three alternatives: Tactical, Team Management, and Other. The third level, Message Content, consists of more detailed information on a message's content and differs according to Message Subject.

---

[1] MacSHAPA is a research software package that is under development at the University of Illinois at Urbana-Champaign and is distributed by Crew System Ergonomics Information Analysis Center (CSERIAC).

Tactical, Team Management, and Other. The third level, Message Content, consists of more detailed information on a message's content and differs according to Message Subject.

Message Type. Message type identifies the purpose for sending a message. A team member either requests information from another team member, replies to a previous message, or provides information to another team member without being asked.

Request. Requests are messages in which one team member asks another team member for information. No restrictions are imposed on what type of information is requested or whether the request can be answered immediately. The frequent use of this category indicates how well team members recognize gaps in their own information. On the other hand, lessfrequent use of this category may indicate that team members provide needed information to other team members without prompting.

Provide. One of the primary tasks facing each team member in the distributed information processing environment is providing other team members with information they need to understand the developing tactical situation. Particular attention should be paid to what information was shared, with whom it was shared, when it was shared, and what events triggered sharing. In this category, in contrast to a request for information, a team member informs another team member based on an estimate of the usefulness of the information to the recipient.

Reply. A reply is a response to a previous request for information or some other communication by another team member on the same topic. It is used most frequently during a running exchange between team members on a single topic and, as such, provides a measure of the number of communication exchanges on a specific topic.

Message Subject. The three Message Subjects—Tactical, Team Management, and Other—are broad enough to describe the full range of communications that occur among team members in a scenario-driven simulation study. Each subject is further subdivided into a set of Message Content categories that specify the nature of the transmitted information. The next section presents the content categories for each subject category.

Tactical Information. Tactical information describes or pertains to specific situational elements integral in developing an accurate and complete tactical picture. Discussing tactical information is one the most fundamental activities in command and control. Tactical information can consist of several types of information, including unaltered raw information, transformed raw information, inferences regarding future events, or alternative courses of action based on raw, transformed, or inferred information.

Team Management Information. Team management information organizes and guides team members' actions toward system goals as well as their interactions in performing various tasks. This information can be used to assign duties and responsibilities to team members as well as to clarify points of confusion and correct misunderstandings among team members. Team members usually need team management information when they are unfamiliar with assigned tasks or inexperienced in working together as a team.

Table 2. Categories used to codify team information processing.

Message Subject/ Message Content	Message Type		
	Request	Reply	Provide
TACTICAL			
Raw Information			
Correlate			
Fuse			
Assess			
Acknowledge			
Clarify			
Recommend			
TEAM MANAGEMENT			
Advise/Direct			
Resolve/Correct			
Structure			
Mediate			
Acknowledge			
Clarify			
OTHER			
General Instructions			
Computer-related			
Social/Unrelated			
Tools/Procedures			
Acknowledge			
Clarify			

Other Information. Team member comments on subjects other than tactical or team management information are placed in the Other category. This category is used to code information that is indirectly related or unrelated to developing an understanding of the tactical situation or to dealing with team management issues. This category ensures that all team member comments are characterized, regardless of their immediate relevance to the stated goals of the assigned task. However, this category is not a repository for irrelevant comments, but a category that provides information that is not classified elsewhere.

Message Content. The message content categories detail several aspects of a message, including the nature of the transmitted information, the goal of the message, or the type of cognitive activity required to generate the message. Because these categories detail each messages' nature, each category is discussed for each message subject.

Tactical Message Content Categories. The following seven categories for tactical message content categorize information that contributes to an understanding of the tactical situation and that identifies information that should be included in a SITREP. Correlate, fuse, and assess represent increasing degrees of preprocessing of raw information before a message is sent.

1. *Raw Information.* Raw information is a request for, reply to, or provision of information that was not transformed or otherwise altered from its original form. In most cases, there are no restrictions on format; raw information may be verbal, textual, graphical, or any combination of these formats. The subset selected from the total pool of raw information demonstrates the focus of team members' attention. Analysis of selected information indicates team members' concerns and hypotheses at different times during an exercise.

2. *Correlate.* Correlate is the process of linking two information items to increase understanding of the relationships among distinct aspects of a scenario. These information items may have been presented at different times and, perhaps, were initially associated with different aspects of the scenario. The correlation formed between them results from cognitively transforming the data rather than simply combining two items that were presented at the same time and place.

3. *Fuse.* Fuse is the process of generating an inference based on existing information about a past, present, or future scenario. The information on which the inference is based can be any combination of raw or processed information. The inference serves as new information that can be tested for accuracy against currently available as well as incoming information. As a consequence of this testing process, the inference may be revised or updated to make it consistent with ongoing scenario events and new information.

4. *Assess.* Assess is the process of evaluating the current situation in terms of mission objectives and the possibility that those objectives can be thwarted. This evaluation can include the operational capability of friendly forces and the threat capability of hostile or suspect forces. Team members may also perform an evaluation of the future threat capability of hostile or suspect forces based on anticipated scenario events.

5. *Acknowledge.* Acknowledge is a process of explicitly declaring the receipt of a message. The routine repeat used in military communications constitutes typical acknowledgment. Ideally, for more complicated messages, an acknowledgment permits the message sender to determine whether the recipient has understood the message's meaning.

6. *Clarify.* Clarify is a process of reiterating or rephrasing a message so that its meaning becomes clear to recipients. Reiteration involves simply resending a message using the same form and content as the original transmission, whereas rephrasing involves changing message form and/or content to enhance its meaning. For example, rephrasing may add information to a message so that its context and purpose become easier to understand.

7. *Recommend.* Recommend is a process of proposing a specific action to support friendly forces, to facilitate the performance of certain tasks, or to counteract enemy threat capabilities.

Team Management Message Content Categories. The six categories that follow were used for categorizing the content of messages about coordinating and directing the team's operations:

1. *Advise/Direct.* The Advise/Direct category is used whenever one team member supplies another team member with information regarding responsibilities for collecting, compiling, and distributing information. This is commonly done in response to scenario events that one or more team members failed to detect, or to remind team members of an approaching deadline for some scheduled activity, such as filing interim SITREPs.

2. *Resolve/Correct.* The Resolve/Correct category involves maintaining the duties and responsibilities of team members' roles. Resolve settles potential points of conflict between team members, often regarding the assignment of responsibility for supplying team members with information. Correcting involves reassigning responsibilities and duties to appropriate team members after they were mistakenly assumed by the wrong individual.

3. *Structure.* The Structure category organizes the duties and communication between individual team members. In this sense, structure is similar to the Advise/Direct category, but it is more general in terms of team member actions.

4. *Mediate.* The Mediate category involves intervening between team members who are encountering difficulty in any of several areas, such as communicating with each other, or in task assignments.

5. *Acknowledge.* The Acknowledge category is used as described in the tactical message content section, but in the context of team management issues.

6. *Clarify.* The Clarify category is used as described in the tactical message content section, but in the context of team management issues.

Other Message Content Categories. The six categories for the Other Message Content categorizes information that pertains to aspects of team member and team performance that were not captured by the Tactical and Team Management message content categories.

1. *General Instructions.* The General Instructions category is for comments regarding the organization, content, procedures, and conduct of the study. These messages often consist of requests for restatements of experimental protocol that was previously provided during the initial training period.

2. *Computer-related.* The Computer-related category is for comments regarding computer operation, such as the use of various interface features, as well as computer malfunctions that may arise during the study. Difficulties encountered while using software applications were included in the Tools/Procedures category (see item 4).

3. *Social/Unrelated.* The Social/Unrelated category is used for comments that are extraneous to the performance of the study task, but which reflect the dynamics of group interaction. This information can prove helpful in interpreting the actions and decisions made by the team, such as the informal chain of command.

4. *Tools/Procedures.* The Tools/Procedures category is reserved for comments about specific tools and the procedures for using them. Malfunctions in tool operation and questions about the steps needed to complete a certain action are noted here.

5. *Acknowledge.* The Acknowledge category is used in the same way as described in the tactical message content section, but in the context of the topics covered by the Other message content categories.

6. *Clarify.* The Clarify category is used as described in the tactical message content section, but in the context of the topics covered by the Other message content categories.

EXAMPLE INVESTIGATION: USE OF SHARED MAP BY C^2W TEAM

The study reported in this section provided an initial assessment of the COLAB as a facility for examining collaborative task performance in a laboratory setting. It also facilitated the development of the analytic techniques reported in the preceding section for evaluating individual and team task performance. This study investigated the use of a shared map by a four-person team performing a situation assessment task that required collaboration among team members to generate an accurate and complete assessment.

SUBJECTS

Five four-person teams of active duty U.S. naval officers participated in a collaborative situation assessment task in the COLAB. The members of each team were experienced in working together as a unit during Fleet operations.

DESIGN

Each team member was assigned one of four operational roles for participating in a scenario-driven situation assessment task: Command and Control Warfare Commander (C^2WC), Cryptological (Crypto) Watch Officer, Electronic Warfare (EW) Watch Officer, or Intelligence (Intel) Watch Officer. When possible, the role assigned to each player corresponded to the duties customarily performed in fleet operations. Tasks associated with each role were analogous to actual operational tasks in terms of the types of information received at each workstation and the necessity for team members to share information in order to understand the tactical situation and generate an accurate SITREP. The EW workstation displayed lines of bearing and alerts containing information on electronic signatures. The Intel workstation provided a steady stream of intelligence data with some source identifications. The Crypto workstation provided information from intercepted communications during the mission. The C^2WC workstation presented alerts and mission tasking information from mission planners and other simulated support sources.

Besides receiving information from the tactical display that was identical to that received by all other team members, each team member also received alert messages keyed to the team member's workstation. The EW Watch Officer also received electronic support measures (ESM) lines of bearing in addition to a unique set of alert messages.

The study scenario consisted of four vignettes, each building upon events in the preceding vignette (Quinn & Gwynne, 1996). The mission's purpose was to conduct a noncombatant evacuation operation (NEO) to remove American citizens from a foreign country where a civil war had started. The vignettes were constructed so that the workload of each team member remained relatively constant across vignettes. In addition, the workload of each team member was approximately the same within each vignette, except for the EW Watch Officer, who received more alert messages than the other positions.

There were two independent variables in this study: Voice and Collaborative Tool. Voice had two levels: With Voice, which featured full voice communications for all team members, and No Voice, which eliminated this capability. Collaborative Tool also had two levels: Non-Shared Whiteboard and Shared Whiteboard. The Non-Shared Whiteboard condition consisted of e-mail and a map of the tactical situation on which each team member could place drawings and text visible to himself but not to other team members. The Shared Whiteboard condition differed from the Non-Shared Whiteboard condition in that the tactical situation map was shared among all team members.

Markings made on it by any team member were visible to every other team member in real time. Each team received all four combinations of Voice and Collaborative Tool conditions, one during each of the four vignettes. The presentation order of these combinations was counterbalanced across teams to control for potential sequence effects.

The purpose of using these four combinations of independent variables was to investigate the effect of different degrees of collaboration on the ability of team members to exchange information with each other to formulate an accurate, complete, and consistent situation assessment. The tools used in the various combinations differed in the extent to which they permitted collaboration. For example, Voice allowed instantaneous information exchange, but left no permanent record for later reference. The Shared Whiteboard, in contrast, required more time and effort to use, but provided a means for communicating graphics and text.

The electronic whiteboard was the major collaborative technology variable. It was included in the study to identify the source of any performance differences across the Shared Whiteboard and Non-Shared Whiteboard conditions. The Shared Whiteboard was a unique communications modality that permitted communications to occur using text and/or graphics. In addition, the Shared Whiteboard permitted the investigation of the concept of shared mental models. This concept was proposed as a way to understand teamwork among individuals working on a common problem in a distributed environment. The degree to which teammates have the same understanding of the tasks to be performed and the resources available for performing those tasks (i.e., a shared mental model) determines the effectiveness of their interactions. In this study, players used the electronic whiteboard to express and develop their shared mental model of the evolving tactical situation.

Previous research found that the presence or absence of voice communication accounts for large differences in many task performance measures. This finding is not surprising because voice communication is a powerful, easily used capability that frequently overshadows other communications modalities. Because voice communication can potentially dominate other modalities, it was not used in two of the four experimental conditions to encourage use of the other communications tools. Voice communication also provided a validation of study procedures by obtaining the same significant advantage for voice that other studies reported.

The factorial combination of the Voice and Collaborative Tool variables permitted interactions between them to be hypothesized. These hypothesized interactions are summarized in figure 2. A main effect for Voice was predicted, in which performance in the Voice condition would be superior to the No Voice condition for both the Shared Whiteboard and Non-Shared Whiteboard conditions. An interaction was hypothesized, in which the No Voice/Shared Whiteboard condition would produce performance approaching that observed for the two Voice conditions due to the added capability for information exchange provided by the Shared Whiteboard. The Voice/Non-Shared Whiteboard condition was expected to produce performance that was marginally lower than the Voice/Shared Whiteboard condition, whereas performance in the No Voice/Non-Shared Whiteboard condition was expected to be significantly lower than in either of the voice conditions or the No Voice/Shared Whiteboard condition.

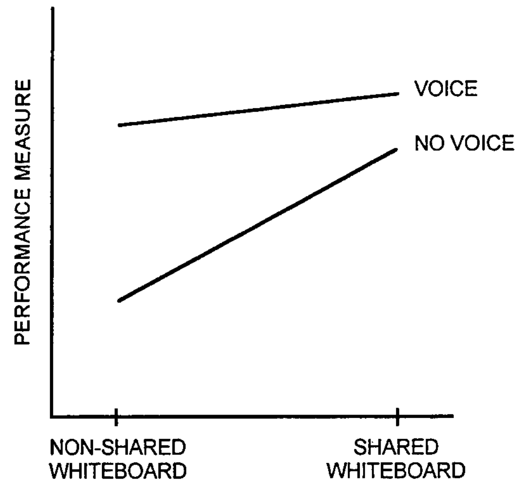


Figure 2. Hypothesized interaction between the Voice and Collaborative Tool variables.

PROCEDURE

A tactical C^2W situation assessment scenario was presented to each team. The scenario was composed of the four vignettes listed in table 3, which also shows the C^2W team objectives for each vignette.

Table 3. Scenario vignettes and team objectives.

Vignette	Vignette Title	C^2W Team Objectives
1	Transit to Amphibious Operations Area (AOA)	Assess hostile country capabilities and intentions
2	Establish AOA-Launch	Prepare plans in support of the operation
3	Conduct Mission	Monitor and support the operation
4	Return to Task Force	Protect the Force

Before beginning the experiment, COLAB personnel briefed team members regarding the nature of the scenario and their roles in it. They were were told that during each vignette, each of them would receive unique information about the tactical situation in alert messages, and that to generate a complete and accurate situation assessment, they must share at least some of those alerts with each other. Each team member was given a set of documents that described the scenario and provided the background information needed to understand and perform the situation assessment task (Quinn & Gwynne, 1996).

Next, COLAB personnel introduced team members to the experimental setting and trained them on how to use the software tools proficiently. A multiphase approach to training was used in which a trainer first demonstrated tool use to team members. Training topics included sending and receiving e-mail, copying and marking alert messages, using the shared map to exchange information, and using the tactical display to acquire information about the changing tactical situation. For team members with little computer experience, instruction was also provided in the basic aspects of using

a Macintosh interface, including use of a mouse, moving from one application to another, scrolling through an active window, and performing basic manipulations of text and graphics (cut, copy, paste). In the second phase of the training program, team members familiarized themselves with the tools during a practice session. During this practice session, they used the tools to perform typical tasks that would occur during the experiment. The trainer provided feedback to ensure that all team members developed proficiency. In the last phase of training, a mastery test was administered to all team members. The test required them to perform basic text and graphic manipulations, send and receive e-mail, mark alert messages, and use the tactical display to identify and characterize individual objects. The goal of this demonstration, practice, and testing procedure was to make each team member self-sufficient in tool use. Although the training was largely successful in achieving this aim, the trainer was always available during the experiment to provide assistance in using the software tools.

Upon completion of training, the four vignettes of the mission were performed, separated by short rest periods. Each team member received a series of alert messages during a vignette. Most of these messages were unique to each team member. Team members were instructed to mark each message as soon as they read it. This permitted each team member to keep track of read and unread messages and to record the time this action occurred for reference during data analysis. If desired, messages could be prioritized according to their perceived significance by multiple marks. The response latency of team members in marking alert messages indicated their attentiveness to the tactical situation. Toward the end of each vignette, each team member completed a SITREP, answering its four questions about the developing tactical situation.

At the conclusion of the experiment, a debriefing was held during which individual team members could provide additional information about their performance in the context of several distinct topics. COLAB peronnel assessed the familiarity of each team member with certain types of human-computer interface features and computer programs, including geographic tactical displays, the Macintosh user interface, word processing, graphics/drawing programs, e-mail/message processing programs and collaboration/groupware programs. Subjects were also asked to rate the utility and usability of tool features for performing certain actions such as pointing to areas on the map, drawing on the map, placing text onto the map, sending messages via e-mail, and copying part of the map and pasting it into an e-mail document. Utility and usability were gauged in terms of how important tool features were for these actions, how well the tools supported team communications, and how easy the tools were to use. Next, subjects were asked to elaborate on their responses to items on the SITREP, providing clarification and greater detail. Finally, subjects were asked to comment on six topics related to their performance of the experimental scenario: verbal communications, situation assessment, data communications, compensatory behavior, team leadership, and communication similarities. Subject responses to questions on each of these topics centered on the role of different communications tools and how the capabilities and limitations of those tools affected their responses to certain tactical situations during the scenario and to the development of their situation assessments.

SITREPs

The SITREPs solicited information from each team member regarding threat force disposition and capabilities, anticipated monitoring activities, and recommended changes to the tactical plan. Figures 3, 4, and 5 display, respectively, the findings for the three measures used to evaluate the SITREPs: Quality, Completeness, and Consistency.

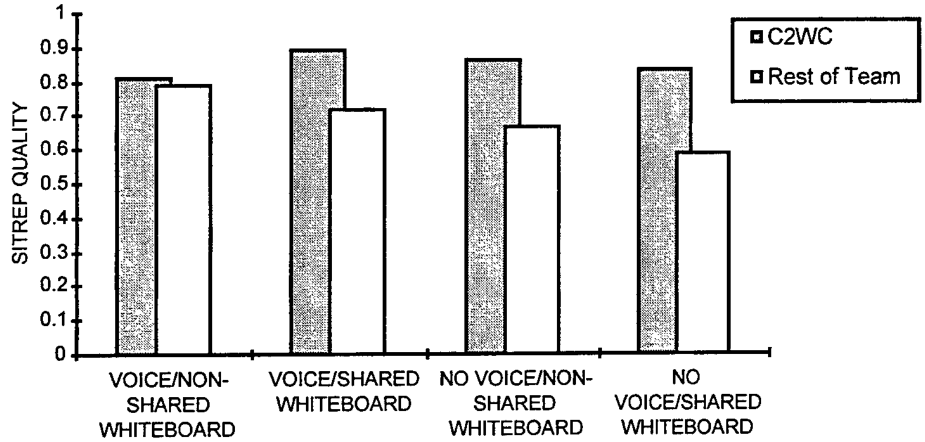


Figure 3. Quality scores for C²WC and rest of team across combined experimental conditions.

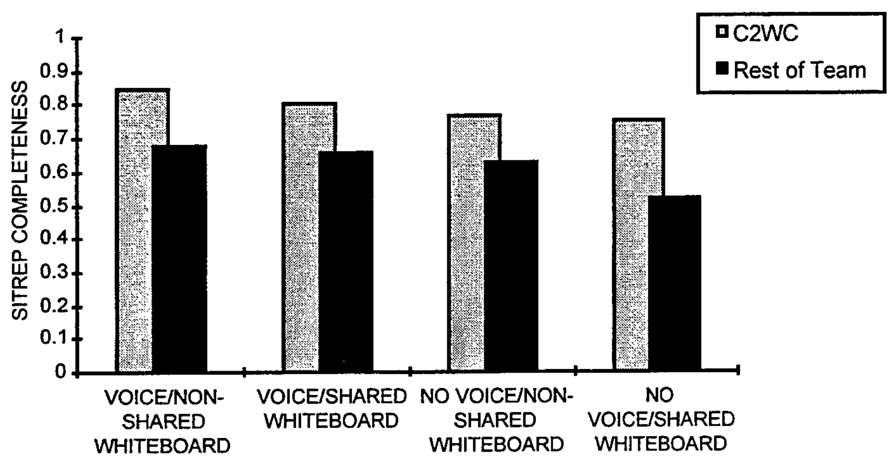


Figure 4. Completeness scores for C²WC and rest of team across combined experimental conditions.

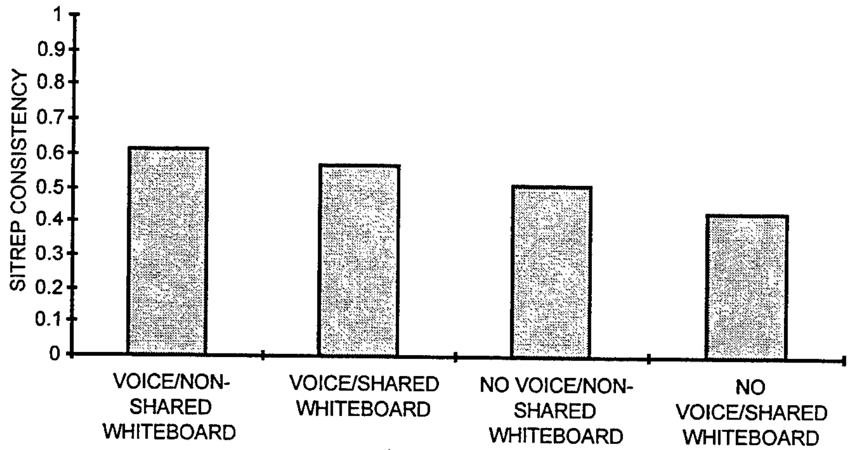


Figure 5. Consistency scores across combined experimental conditions.

Figures 3 and 4 show that the C^2WC obtained higher Quality and Completeness scores than the rest of the team in every experimental condition. This finding suggests that relevant information was being passed from team members to the C^2WC more than it was being passed among themselves or from the C^2WC to team members. The two Shared Whiteboard conditions were associated with lower scores than the two Non-Shared Whiteboard conditions on all three measures. In addition, the No Voice condition was associated with lower scores than the Voice condition. This latter finding suggests that a communications tool capable of handling a high volume of messages (i.e., voice) helped create a greater awareness of threat areas than lower volume tools. However, in most cases, group differences were small and subject to questions concerning their reliability given the small sample size (n = 5) in this study. The lack of a clearly interpretable trend in the differences among the four experimental conditions suggests that each condition provided subjects with adequate means to transmit tactical information to the C^2WC.

Figures 6, 7, and 8 display these findings for the two levels of each independent variable on the Quality, Completeness, and Consistency measures, respectively. The Voice scores are pooled across the Whiteboard variable and the Whiteboard scores are pooled across the Voice variable. Once again, no clear pattern emerges, aside from the superiority of the C^2WC on the Quality and Completeness measures.

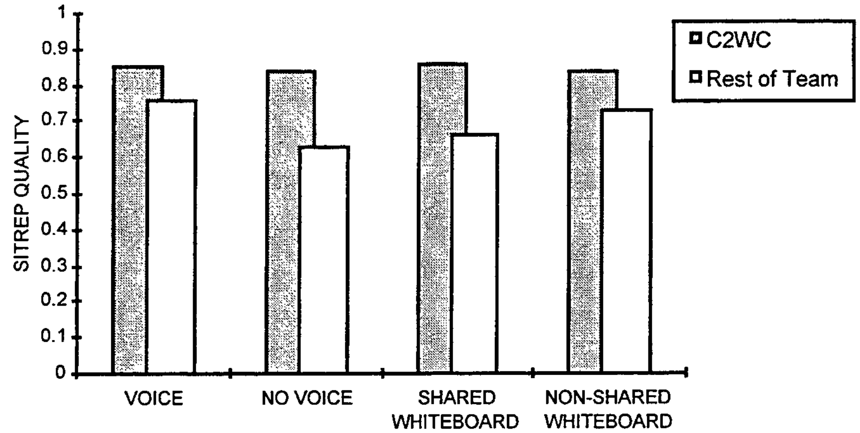


Figure 6. Quality scores for C^2WC and rest of team across separate experimental conditions.

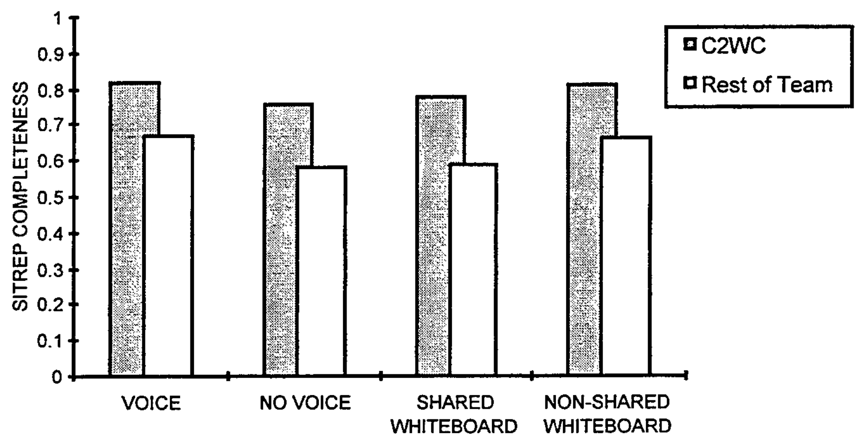


Figure 7. Completeness scores for C^2WC and rest of team across separate experimental conditions.

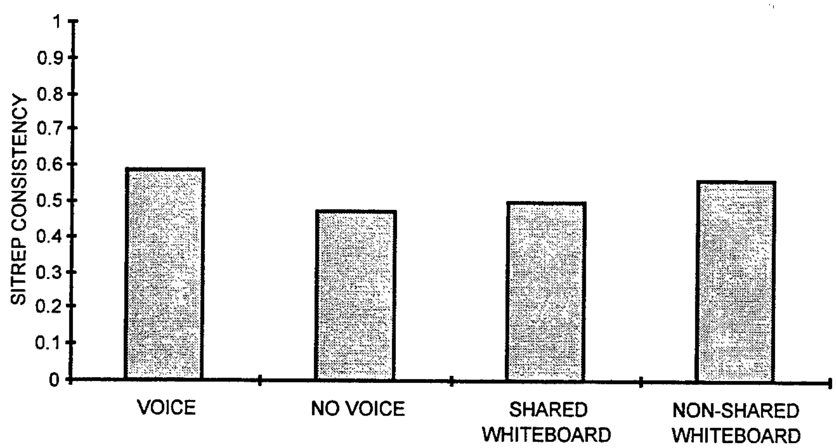


Figure 8. Consistency scores across separate experimental conditions.

MARKING OF ALERT MESSAGES

The percentage of marked alert messages indicates the extent that team members complied with experimenter instructions. Table 4 shows that most team members complied with the experimenter's instructions to read and mark the alert messages. However, a few consistently failed to comply with this request, as indicated by the shaded cells in table 4. In particular, two team members in the USN 3 team marked only a small number of alert messages in all four vignettes and one team member in the USN 4 team did likewise in two of the four vignettes. The failure of these team members to mark most of the alert messages in more than one vignette suggests that their noncompliance was not due to excessive task demands, but rather to lack of compliance with or inattentiveness the experimenter's instructions.

Table 5 presents the mean number and percentage of alert messages that were marked in terms of experimental condition. The percentage marked ranged from about 50% to nearly 100%. This variability did not appear to be associated with experimental conditions; instead, it appeared to be due mainly to the failure of certain team members (those shaded in table 4) to comply with experimenter instructions to mark the alert messages.

The response latency to marking alert messages (defined as the time that elapses from the appearance of an alert message to the time when it is first marked) indicates whether team members remained current with the scenario events. As table 6 shows, there was considerable variability in these values; however, no clear pattern emerged with respect to experimental condition or team member role.

If outliers are excluded, the mean for the No Voice/Shared Whiteboard condition becomes 2:07 (min:sec) and that for the No Voice/Non-Shared Whiteboard condition becomes 1:41. While the latter value thus becomes similar to those obtained for the two Voice conditions, the former value still remains higher. This result suggests that the No Voice/Shared Whiteboard condition was associated with longer response latencies than the other three conditions. A plausible explanation for this finding is that the No Voice condition produced more frequent use of e-mail, which was much more laborious to use than Voice. Likewise, the Shared Whiteboard condition could also draw team members' attention away from incoming alert messages.

Table 4. Number of alert messages marked by each team member in each vignette.

	Total Messages Presented	Team				
		USN 1	USN 2	USN 3	USN 4	USN 5
Vignette 1						
C^2WC	7	7	7	2	3	7
Crypto	15	14	15	15	15	15
EW	19	19	19	4	19	7
Intel	12	12	12	10	2	6
Vignette 2						
C^2WC	9	9	9	0	8	9
Crypto	18	12	18	17	18	18
EW	12	12	12	0	11	11
Intel	17	9	17	17	17	15
Vignette 3						
C^2WC	8	8	7	0	0	8
Crypto	10	7	10	10	10	10
EW	30	30	30	0	30	15
Intel	10	10	10	8	3	10
Vignette 4						
C^2WC	9	8	9	0	7	*
Crypto	14	12	14	14	14	*
EW	16	15	16	0	15	*
Intel	12	11	12	12	12	*

* Vignette was terminated prematurely due to equipment malfunction.

Table 5. Mean number and percentage of alert messages marked according to experimental condition and team member role.

Role	Total Messages	Voice/ Shared Whiteboard	Voice/ Non-Shared Whiteboard	No Voice/ Shared Whiteboard	No Voice/ Non-Shared Whiteboard	Mean Number Marked	Mean Percentage Marked
C^2WC	33	27.5	23	18.5	16	21.2	64
Crypto	57	54	56.5	54	57	55.4	97
EW	77	61	40	56.5	61	54.6	71
Intel	51	48	46	40	45	44.8	88
Totals	218	190.5	165.5	169	179		

Table 6. Mean response latencies (min:sec) to alert messages according to experimental condition and team member role.

Role	Voice/ Shared Whiteboard	Voice/ Non-Shared Whiteboard	No Voice/ Shared Whiteboard	No Voice/ Non-Shared Whiteboard	Mean
C^2WC	2:50	1:42	1:08	1:15	1:38
Crypto	1:38	1:52	2:02	2:17	1:57
EW	0:52	1:25	3:16	2:43	1:59
Intel	1:14	1:23	2:14	1:07	1:30
Mean	1:38	1:24	2:10	1:51	

Table 7 summarizes the preceding tables by showing the mean response latencies and percentage of alert messages marked according to experimental condition. It appears that the mean latencies are longer in the No Voice condition than in the Voice condition. However, there was one outlier in each of the two Collaborative Tool conditions under No Voice that increased the mean latencies. Recomputing the latencies after eliminating these outliers produced values much closer to those obtained under the Voice condition. The results for the mean percentage of alert messages marked were likewise inconclusive.

Table 7. Mean response latencies and percentage of alert messages marked according to experimental condition.

Experimental Condition		Mean Response Latencies (mm:ss)	Mean Percentage Marked
Voice	Non-shared Whiteboard	1:42	76%
	Shared Whiteboard	1:38	87%
No Voice	Non-shared Whiteboard	2:57[1]	82%
	Shared Whiteboard	2:47[2]	78%

[1] Elimination of outlier (8:34) yields mean of 1:41
[2] Elimination of outlier (9:50) yields mean of 2:07

TIMELINE ANALYSIS OF TEAM PERFORMANCE

An exploratory sequential data analysis technique (MacSHAPA) was applied to the verbal, e-mail and Whiteboard communications protocol data. The sequential analysis capabilities of MacSHAPA permit temporal relationships among variables to be explored more thoroughly than is possible with more traditional techniques. The major purpose of this analysis was to evaluate the relationships among scripted scenario events, such as the presentation of alert messages, and team member communications and actions in response to those events.

The following analyses compared teams across experimental conditions and/or experimental conditions across teams:

1. Proportions of communications by mode (e.g., voice, e-mail, Whiteboard);
2. Proportions of communications by type (e.g., request, reply, provide);
3. Proportions of communications by content (e.g., raw, correlate, fuse);
4. Proportions of communications by threat areas.

A finer level analysis involved team members within teams by experimental conditions:

1. Numbers and proportions of alert messages acknowledged;
2. Latencies between system initiation of alert message and team member acknowledgment;
3. Types of communications initiated.

The most detailed level of analysis investigated individual team members regarding:

1. Intended recipient(s) of specific types of communications by condition;
2. Interactions regarding a specific threat area by condition.

To illustrate the type of analyses that were performed, Figure 9 shows player actions that were generated in two separate sequences. The first sequence involved player response to an alert message. The message was received by Crypto approximately 20 minutes into the vignette. Shortly thereafter, Crypto acknowledged the message and then informed team members by voice that hewould send them a message. Crypto then e-mailed the message to all team members, and each team member acknowledged receipt by voice. A new sequence of actions then began with Intel calling EW regarding a certain threat area and EW acknowledging the communication. Intel then assessed the extent of the threat, which was acknowledged by EW.

SUMMARY OF TOOL USE

Three primary tools were used by subjects in this study to transmit information among themselves: verbal communication, e-mail, and a whiteboard that displayed a snapshot of the tactical situation taken periodically during a vignette. The preferential use of these tools under different experimental

conditions clarify the factors that determine the selection of a particular tool and the type of communication that occurs with that tool. When these factors are understood, they can be used to develop, test, and evaluate interfaces for collaborative applications. Figure 10 summarizes the number of messages sent as a function of tool used and experimental condition. Although there was an increase in the use of e-mail and the whiteboard for communications, it is apparent that there was a large decrease in the number of messages transmitted in the No Voice conditions relative to the Voice conditions. The fact that there was no overall difference in the quality, completeness, or consistency of the SITREPs associated with the No Voice conditions indicates that the additional messages transmitted by voice did not appreciably enhance the players' understanding of the tactical situation. This conclusion assumes that the Quality, Completeness, and Consistency measures were sufficiently sensitive to detect meaningful differences, an assumption still unverified.

Verbalizations

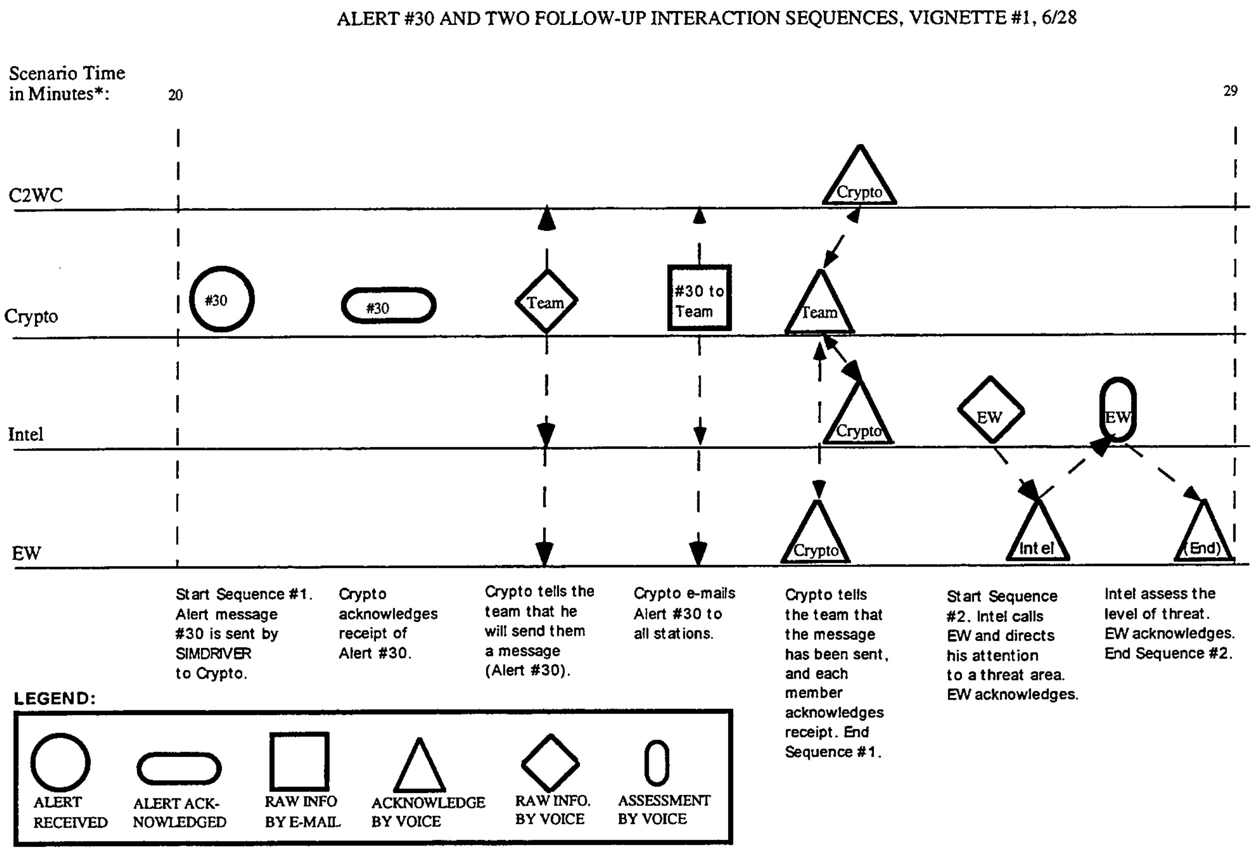
An open communications net was used in vignettes that permitted verbal communications between team members. All team member comments were recorded on audiotape and transcribed. The verbal transcriptions were coded and the geographic area of interest was noted to determine the focus of each team member's attention at various times during a vignette.

Frequency of Verbalizations. The number of verbalizations did not differ significantly as to whether the Shared Whiteboard or Non-Shared Whiteboard conditions were used (985 verbalizations in the Shared Whiteboard condition vs. 1029 in the Non-Shared Whiteboard condition). As table 8 shows, a similar pattern of results was obtained when verbalizations were compared in terms of the major message types (Request, Reply, Provide) and message subjects (Tactical, Team Management, Other).

Type of Information Transmitted by Verbalizations. The upper part of table 8 shows that the majority (58%) of verbalizations were replies to prior communications. This finding is understandable given the scoring method for verbalizations that involved categorizing a message as a reply when it was either a response to an immediately preceding request for information, a provision of information, or a response to a running exchange between two or more players on the same topic. Providing information (22%) and requesting information (19%) were less common than replies, and the three message types did not differ between the Non-Shared Whiteboard and Shared Whiteboard conditions.

Turning to Message Subject, the lower part of the table shows that most verbalizations pertained to tactical information (82%), with the Other and Team Management categories seldom used. As with Message Type, there were no meaningful differences between the Non-Shared Whiteboard and Shared Whiteboard conditions.

Relation of Verbalizations to Type of Tools Used. The reason for the lack of differences between the Non-Shared Whiteboard and Shared Whiteboard conditions is likely related to the dominant role of verbalizations in team member communication. Although other tools were freely available in both the Non-Shared Whiteboard and Shared Whiteboard conditions, most information was transmitted verbally. Future studies could probe more deeply into the relationship between verbalizations and the type of tools used by degrading the quality of verbal communications so that they are not as reliable and freely available as they were in this study.



**Note: interactions shown above occurred within a 9-minute time segment of the scenario and are reflective of the order of occurrences; however, to facilitate the graphical display of these interaction sequences, time intervals between them have been modified.*

Figure 9. Example of sequential interaction analysis of team member responses.

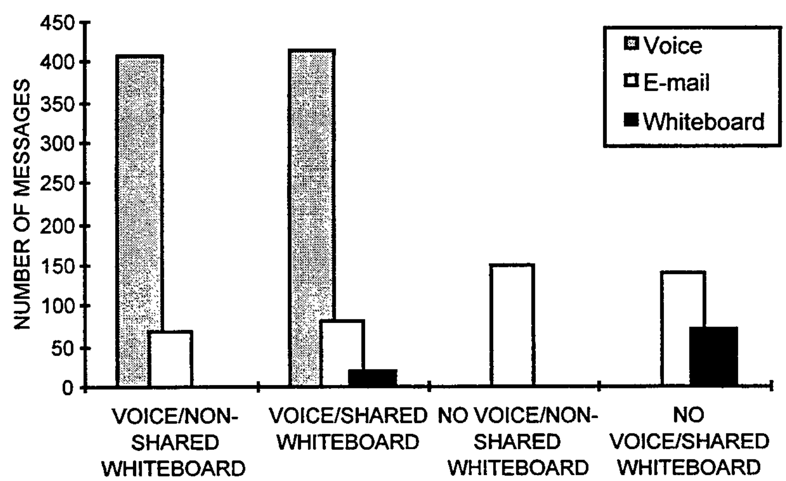


Figure 10. Number of messages transmitted among team members using different tools, by experimental condition.

Summary of Verbalization Findings. As expected, verbal communications were direct, immediate, and powerful in passing information among team members. When verbal communication was available, it became the primary means for transmitting information. The ease of use of its interface, coupled with the immediacy of verbal transmissions, resulted in its constant, preferential use whenever it was available.

One drawback of verbal communications is the lack of graphics, which can be crucial to understanding a geographically oriented task such as C^2W. Usually, team members recognized this drawback. In cases where verbalizations alone were inadequate to convey the full meaning of a message, team members could supplement them with drawings if they were in the Shared Whiteboard condition, or by attaching a graphic as an e-mail enclosure if they were in the Non-Shared Whiteboard condition.

A second drawback to verbal communications is that there is no permanent record. Instead, the message recipient must understand the communications correctly on first hearing and remember them accurately. Under conditions of high mental workload, however, human memory may prove inadequate to accurately recall prior verbal messages. E-mail and Whiteboard messages, to be discussed below, can provide textual or pictorial memory cues that reduce memory-related performance deficits. Team members had more difficulty recognizing the problems associated with a lack of a permanent record than they did recognizing the importance of graphical enhancements to a message.

E-mail

E-mail was available to all team members in all experimental conditions. In contrast to verbalizations, e-mail messages provided a permanent record so that message recipients could review them any time after receiving them. An e-mail message could be directed to a single team member or any combination of team members. The typical range of graphical user interface manipulations were available in e-mail. For example, alert messages could be pasted into an e-mail

Table 8. Proportion of verbalizations in major message categories for Shared Whiteboard and Non-Shared Whiteboard conditions.

Category	Non-Shared Whiteboard (n=1029)	Shared Whiteboard (n=985)	Overall (n=2014)
Message Type			
Request	0.18	0.20	0.19
Reply	0.60	0.57	0.58
Provide	0.22	0.23	0.22
Message Subject			
Tactical	0.81	0.84	0.82
Team Management	0.05	0.02	0.04
Other	0.13	0.14	0.13

message and a snapshot (screen capture image) could be enclosed with a message. However, these capabilities were associated with an interface that team members found the most difficult and time-

consuming of all the tool interfaces to learn and use, especially when attaching an enclosure to a message.

Frequency of E-mail Messages. Predictably, team members used e-mail more frequently when verbal communications were unavailable. This was true in both the Non-Shared Whiteboard and Shared Whiteboard conditions. Table 9 shows that in the Non-Shared Whiteboard condition, 239 e-mail messages were sent. Of these, 152 were sent by team members in the No Voice condition, compared to 87 in the Voice condition. In the Shared Whiteboard condition, team members sent 242 e-mail messages, 160 while in the No Voice condition and 82 in the Voice condition.

Type of Information Transmitted by E-mail Messages. Message type categories (Request, Reply, Provide) in table 9 show that providing information was the main purpose of e-mail messages, accounting for an identical 74% of the total e-mail messages in both the Non-Shared Whiteboard and Shared Whiteboard conditions. Likewise, replying to prior messages accounted for 17% of all e-mail messages, while requesting information accounted for 9% in both toolset conditions. There were differences between the Voice and No Voice conditions because relatively more messages fell into the Request and Reply categories in the No-Voice condition; consequently, a lower percentage of e-mail messages provided information in the No Voice condition. This pattern occurred in both the Non-Shared Whiteboard and Shared Whiteboard conditions.

Turning now to Message Subject, a similar pattern of results was obtained across conditions, although relatively more messages pertained to topics in the Other category in the Non-Shared Whiteboard condition. Most of these messages concerned computer- and tool-related topics, such as how to attach enclosures to an e-mail message. As with the verbalization results, communications concerning team management topics were rare, an expected finding, given the experience of the team members working with each other.

Table 9. Proportion of e-mail messages in major message categories.

Category	Non-Shared Whiteboard			Shared Whiteboard		
	Voice (n=87)	No Voice (n=152)	Overall (n=239)	Voice (n=82)	No Voice (n=160)	Overall (n=242)
Message Type						
Request	0.05	0.11	0.09	0.04	0.13	0.09
Reply	0.05	0.24	0.17	0.07	0.22	0.17
Provide	0.91	0.65	0.74	0.89	0.66	0.74
Message Subject						
Tactical	0.78	0.89	0.85	0.91	0.94	0.93
Team	0.03	0.01	0.02	0.00	0.04	0.02
Other		0.11	0.13	0.09	0.02	0.04

Whiteboard

The Whiteboard condition consisted of a commercial software package that provided team members with a capability analogous to group use of a Whiteboard. A snapshot of the tactical situation display was taken at periodic intervals during each vignette. This snapshot drawing was displayed in the Whiteboard window; and each team member could make text or line drawing entries on it. In the Shared Whiteboard condition, these entries could be seen by all other team members.

Frequency of Whiteboard Messages. Predictably, Whiteboard messages were more numerous when voice communications were unavailable. Table 10 shows that of the 95 Whiteboard messages, 76 were sent when no voice communications were available. The Voice and No-Voice conditions showed similar patterns of results when messages were compared in terms of major message types (Request, Reply, Provide) and message subjects (Tactical, Team Management, Other). The overall pattern of Whiteboard messages was similar to that of the e-mail messages.

Type of Information Transmitted by Whiteboard Messages. The upper part of table 10 shows that the majority of Whiteboard messages, whether text or graphics, provided information to other team members, as opposed to requesting information or replying to a previous message. The lower part of the table shows that, as with the verbal and e-mail messages, the subject of the Whiteboard messages was almost entirely tactical information; messages concerning team management issues or other information were infrequent. Players consistently sent the same pattern of message subjects regardless of tools used. This finding implies that the nature of the task rather than the tools used to perform the task is the critical factor in determining player actions.

An important point concerns how the Whiteboard was used. Team members often used their unique Whiteboard cursor to point out specific objects, tracks, or geographic regions. Indeed, it is likely that the majority of communications involving the Whiteboard consisted of this type of action. However, pointing leaves no physical indication that any communication occurred. Although this fact implies that the present analysis of Whiteboard use is incomplete, it may be argued that team members made physical entries on Whiteboard for information they considered to be the most important. If true, then the messages analyzed here constitute the most crucial information transmitted by means of the Whiteboard.

Table 10. Proportion of Whiteboard messages in major message categories for Voice and No Voice conditions.

Category	Voice (n=19)	No Voice (n=76)	Overall (n=95)
Message Type			
Request	0.16	0.16	0.16
Reply	0.00	0.04	0.03
Provide	0.84	0.80	0.81
Message Subject			
Tactical	1.00	0.95	0.96
Team Management	0.00	0.05	0.04
Other	0.00	0.00	0.00

Team members tended to use the Whiteboard to reiterate, emphasize, or clarify messages previously transmitted using a different tool. Less frequently, the Whiteboard was used as the first tool to transmit a message. At times, pictorial and textual information would be combined into one message, as when an alert message was pasted onto the Whiteboard and an arrow or circle used to highlight the geographic area of concern in the message.

POST-EXPERIMENT DEBRIEFING

All team members were debriefed following completion of the experiment to obtain their comments on the effect of prior computer experience on their performance, and the usability and utility of the collaborative tools.

Role of Prior Computer Experience on Task Performance

Table 11 summarizes team member responses to a questionnaire assessing computer program use. Team members had prior experience with most of the types of computer programs used in the experiment. They had the least experience with graphics/drawing programs and collaboration/groupware programs such as Whiteboard. Inexperience with collaboration programs could have caused them to rely more on more familiar tools, such as voice and e-mail. Consequently, this study may have measured the extent to which subjects were willing to experiment with an unfamiliar program along with the utility of collaboration/groupware programs for C^2W tasks.

Table 11. Amount of subject experience with various types of computer programs (percentage of subjects).

Computer Program Type	No Prior Experience	Aware of Such Programs	Use Such Programs Periodically	Use Such Programs Frequently
Geographical tactical display	0	15	20	65
Macintosh user interface	15	10	35	40
Word processing	0	5	35	60
Graphics/Drawing programs	25	30	35	10
E-mail/Message processing programs	0	15	40	45
Collaboration/Groupware programs	15	30	40	15

Usability/Utility of Tools

Table 12 summarizes team member responses to a questionnaire regarding the usability and utility of the tools used in the experiment for accomplishing certain actions. Subjects were asked to rate five frequently performed actions in terms of three criteria: importance for performing a task, support of team communication, and ease of use. A five-point rating scale was used, with greater tool capability being indicated by higher values and vice versa. In general, subjects found all the tools to be at least moderately useful and able to support the tasks required by the experimental scenario. Tools that received somewhat lower ratings, such as Whiteboard, were those with which the subjects were least familiar. This finding suggests that there was an interaction between a subject's experience with a tool and its judged usefulness.

Table 12. Usability/utility ratings of software tool capabilities.

Question\Response	Mean	Standard Deviation
How important were the following capabilities for performing your task?		
Pointing to areas on the map	4.00	0.72
Drawing on the map	3.95	0.89
Placing text onto the map	3.95	0.94
Sending Messages via e-mail	4.40	0.75
Copying parts of the map and pasting it into an e-mail document	3.35	1.23
How well did the following capabilities support team communication?		
Pointing to area on the map	3.95	1.00
Drawing on the map	4.15	0.88
Placing text onto the map	4.20	0.77
Sending messages via e-mail	4.40	0.88
Copying parts of the map and pasting it into an e-mail document	3.70	0.98
How easy to use were the following capabilities?		
Pointing to areas on the map	3.95	0.76
Drawing on the map	3.65	0.99
Placing text onto the map	3.90	0.91
Sending messages via e-mail	3.75	0.97
Copying part of the map and pasting it into an e-mail document	2.85	1.39

CONCLUSIONS

This initial investigation demonstrates the viability of the COLAB for the experimental examination of task performance in simulated navy environments requiring distributed information processing. Results indicate that the COLAB provides an effective means for investigating collaborative tool features and military team communication variables under distributed conditions. Additional research is needed, however, to identify additional configurations for the COLAB that promote the transmission of verbal, textual, numerical and graphical information under varying degrees of collaboration.

No consistent pattern of differences was obtained among the experimental conditions, which differed according to the type and amount of collaboration between team members. A plausible explanation for this result may lie in the nature of the tasks used in this study and the relationship of those tasks to the tactical information provided to each team member. Future studies should vary task characteristics and the information that is provided to team members in order to elucidate the aspects of collaborative tools that facilitate navy C^2W activities.

Collaborative Performance

Team performance was evaluated primarily according to the responses on the SITREP. In all cases, the C^2WC scored higher than the team on the Quality and Completeness measures.

This finding indicates that the team transmitted information to the C^2WC that enabled the C^2WC to create a better situation assessment than the rest of the team. The lack of significant differences on the C^2WC Quality and Consistency scores across experimental conditions indicates that all tool combinations enabled this information transfer. Alternatively, the Quality and Consistency measures

may not have been sensitive enough to detect differences that actually existed. Further research on these measures may suggest alternative approaches to evaluating the SITREPs.

Given research suggesting the value of the features of collaborative tools (Cascio, 1995; Linville, Liebhaber, Obermayer, & Fallesen, 1989; Losada & Markovitch, 1990; McGrath & Hollingshead, 1994), additional studies are needed to isolate the specific effects of various tool features on task performance. A function and cognitive task analysis could identify mental workload demands, processing capacity requirements, and the memory load imposed by a given task. This information could then be used to design studies that examine the suitability of specific tool features for C^2W tasks.

Differences in the number of messages sent using various communications tools reflected either the ease of use of a tool or restrictions in the available tools. Subjects paid less attention to the product generated by a given tool than to the practical aspects of sending a message using that tool. Thus, in certain situations where a permanent record would have been desirable (e.g., when the information referred to a potential future event), a tool was used even though it provided no record (e.g., voice). One solution to this shortcoming would be to provide a recording and retrieval capability for all tools.

Usually, subjects relied on more familiar tools, such as e-mail and voice communications, even in those instances where the Whiteboard capability would have supplied a more immediate or permanent means of information transmission. Most subjects' relative inexperience with collaborative software may have caused them to use the Whiteboard capability less frequently than more experienced subjects. Additional training may be needed to encourage subjects to use the Whiteboard tool routinely.

Cognitive Task Analysis

A cognitive task analysis can be helpful in understanding the factors that contribute to task performance and how they may be influenced by tool features. In addition to specifying the procedures, practices, equipment, and supplies needed to perform a task, a cognitive task analysis can include assessments of mental workload, processing capacity demands, cognitive effort, and attention allocation requirements. This information can direct the selection of the experimental tasks and the tool features that are examined in future COLAB studies.

A preliminary cognitive task analysis conducted as part of this project suggests that numerous C^2W operational tasks have potential sharing and collaboration requirements; the performance of these tasks may therefore be facilitated by collaborative tools. A partial listing of these tasks includes:

1. *Communications: Message and Text Preparation and Recording.* This task involves using all non-voice communications media to transmit information among team members, including text, numerical, and pictorial media. A permanent record is created in the process of preparing each of these message types, providing a means for future reference.

2. *Communications: Voice Transmission and Recording.* This task involves using verbal communications to transmit information from one team member to another. No working record is available when using verbal communications.

3. *Pattern recognition.* This task involves discerning configurations among a set of information elements that are diagnostic for certain situations and conditions. Once a pattern has been

recognized, other activities in C^2W operations such as information interpretation/correlation and situation assessment can occur.

4. *Information interpretation/correlation.* This task involves the processing of raw information to some higher level form; for example, processing may involve correlating or fusing multiple items of information to generate new inferences, threat assessments, or recommendations for courses of action.

5. *Situation assessment.* This task entails compiling information from multiple sources to create an assessment of a dynamically changing tactical situation, and to develop recommendations based on that assessment for pursuing certain courses of action.

6. *Course of action analysis.* This task entails evaluating various alternatives to reaching a stated mission objective, often with the assignment of a likelihood estimate.

SUMMARY AND RECOMMENDATIONS FOR FUTURE RESEARCH

Major accomplishments of this project include the establishment of a research laboratory that enables collaborative activities to be investigated systematically and the development of analytical and statistical techniques for evaluating the performance of team members working on collaborative tasks. In addition, this project led to an improved understanding of the nature of collaboration in navy command and control tasks; this understanding can be applied to future studies of collaboration as well as to numerous other command and control contexts where collaboration plays a significant role in determining system performance and mission success.

ESTABLISHMENT OF COLLABORATIVE RESEARCH FACILITY

A research laboratory was established specifically designed and equipped to investigate collaborative activities among team members as they perform tasks that require them to share information. The laboratory consists of a control room that houses two computers that control the experiment and capture data, and a work room that contains four computers to be used by subjects as they perform duties required by various experimental scenarios. The control and work rooms are linked by digital, video, and voice communications networks that display scenario data, monitor subject performance, and compile and record data produced at each workstation.

Both individual and team performance can be precisely measured as variables are manipulated to determine their singular and combined effects on performance. Data collection capabilities include logging data of all subject actions, human-computer interaction events for all software applications, and time-stamped recording of all communications. All subject actions and verbalizations with respect to the experimental scenario are recorded, documented, and saved for subsequent analysis. The laboratory provides comprehensive experimental control and monitoring as well as thorough records of subject performance.

DEVELOPMENT OF ANALYTIC TECHNIQUES FOR COLLABORATIVE DATA

Data reduction and analytic techniques were developed during this project that can be used for numerous other collaborative study applications in addition to the experiment documented in this report. Taken together, they represent a valuable method for analyzing empirical human performance data collected in collaborative experiments.

The data reduction techniques prepare, condense, and summarize the complex and voluminous protocol data for subsequent statistical analyses. The statistical analyses identify critical subject responses, factors that influence individual and team performance, the interactions among those factors, and how the responses of individual subjects influence the team's collaborative performance.

One analytic technique evaluates the nature and extent of team member collaboration. It derives team performance measures from either verbal or written reports made by each team member in response to scenario-specific questions. Three team performance measures were used in the experiment: Quality, Completeness, and Consistency. Although the exact data used to calculate these measures will vary from one study to the next due to different experimental contexts, the measures themselves are easily adapted to a wide variety of contexts.

Another analytic technique uses timeline and frequency analysis to detect patterns of communication interactions mapped to critical events in the experimental scenario. Interaction sequences among team members are also recorded. The results of this analysis can be plotted for

either individual team members or the entire team to generate a graphical summary of interaction sequences, communication patterns, and information flows used by the team.

The categorization of subject responses in a common category rating scheme allowed subject responses to be analyzed, regardless the type of collaborative tool used to make a response. This permitted verbalizations, e-mail messages, and shared workspace text and graphic messages to be evaluated using the common frame of reference furnished by the rating scheme. This technique facilitated the evaluation of the extent and type of information sharing that occurred among team members.

FUTURE COLLABORATIVE RESEARCH

The results of this study suggest several directions for future collaborative research. A series of mini-experiments could be conducted in which each mini-experiment addresses a specific issue related to the use of collaborative tools. A series of research hypotheses could be formulated to examine users' performance and the effectiveness of various collaborative tools under the given experimental conditions. These experiments could incorporate procedures for (1) identifying collaborative technologies that present the greatest value to navy teams and (2) identifying those functions, tasks, and activities that benefit from collaborative technologies.

Example: Perform a series of mini-experiments that permit collaborative technologies to be evaluated using three tasks common to a broad range of C^4I activities:

1. Mission data collection and organization in C^2W
2. Mission information correlation in C^2W
3. Mission pattern recognition in C^2W

Performance of each of these tasks would be investigated in scenarios previously developed and tested. The effect of the independent variables on performance could be assessed using dependent measures that were developed during the present project. These include C^4I report accuracy, C^4I report completeness, response time, and the user's subjective evaluation of the collaborative feature. Two independent variables that have been established for investigating collaborative features are Mission Type (e.g., Self-Defense, Surveillance, Humanitarian, Planning), which provides an operational context for a scenario, and Tool Features, which include interface and collaborative attributes that have been identified by human factors and military experts as most likely to enhance team performance.

EXTENDING COLLABORATIVE RESEARCH TO OTHER CONTEXTS

During this project, it became apparent that collaboration could be studied in several contexts. One such context is provided by the Decision Support System (DSS) being developed as part of the (TADMUS) project. The DSS is an innovative tool that holds considerable promise as a decisionmaking aid for warfare commanders in tactical command and control settings. Ongoing research has demonstrated that DSS modules can facilitate cognitive aspects of decision making in dynamic threat environments. Further development and refinement of the DSS is needed, however, to maximize the effectiveness of its individual modules as well as to better understand how certain modules interact with each other to influence individual and team performance. This development

and refinement can benefit from the findings and lessons learned from the collaborative study documented in this report.

The Response Manager module of the DSS is especially well-suited to support collaboration among team members. The Response Manager furnishes team members with a prioritized interactive checklist of tactical actions that should be performed to ensure mission success. The prioritization of these actions is displayed via a distance and/or time window that specifies when each action should be performed. The current version of the Response Manager presents a static, representative set of tactical actions that might be performed for air tracks and surface tracks. It would be more desirable, however, for a commander to be able to customize the Response Manager to assess factors that are unique to each mission, such as battle orders, geographic variables, and rules of engagement (ROE). To achieve this goal, the Response Manager will be supplemented with an authoring capability that will permit commanders to augment the standard protocols with information that is highly specific to each mission operation. This authoring capability will be composed of three basic functions:

1. *Authoring.* Authoring will provide the basic text and graphics editing features needed to provide the customized information that is required for making tactical decisions.

2. *Tailoring.* Tailoring will allow modifying the original customized information in response to changing tactical situations and resources. It will provide team members with highly topical and specific information that can expedite the decisionmaking process.

3. *Monitoring.* Monitoring will track the progress of the team by observing the actions of individual team members as they perform tasks using information displayed by the Response Manager. It also involves assessing the quality of the decisions as they are made. If necessary, additional information can be provided as part of an iterative tailoring process.

A Response Manager with an authoring capability can be incorporated at several different levels of team collaboration, including the fleet, battle group, and individual ship. For the purposes of providing an initial demonstration of the value of an authoring capability for the Response Manager, the battle group and ship levels will be emphasized. Subsequent studies will investigate the collaborative issues and requirements involved in implementing Response Manager authoring capabilities at the fleet level.

Findings and issues identified during the collaborative project will be used to guide the development and testing of the Response Manager authoring capability, including:

1. Determine what information should be shown to which team members. Some information may be most appropriately shown only to certain team members but not to those distracted by it. Other information may be relevant for all team members.

2. Achieve a balance between overview information and detailed information. Overview information can direct users' attention to new developments while detailed information can focus users on critical events that deserve close scrutiny. Displays should support users in switching their focus from details and provide overview information that does not interrupt ongoing activities such as threat assessment, hypothesis generation and hypothesis testing.
3. Identify the type and amount of incoming information that may impose excessive mental workload on users. Examine information presentation strategies such as categorization. which can reduce mental workload. Identify cues, such as object categorization schemes and

pattern matching cues, that can support rapid access to long-term memory contents that support the tasks being performed.

4. Determine the appropriate level of abstraction for representing tactical information to team members. This determination should be consistent with the task performance requirements of each team member. The appropriate level of abstraction may vary from one team member to another, depending upon the types of tasks that they are performing.

It is likely that additional collaborative factors may emerge as the authoring capability for the Response Manager is developed, tested and refined. The effects of these factors on collaboration among team members should be investigated concomitantly with the development of the authoring system.

REFERENCES

Cascio, W. R. 1995. "Whither Industrial and Organizational Psychology in a Changing World of Work?" *American Psychologist, 50* (11), 928–939.

Linville, J. M., M. J. Liebhaber, A. H. Obermayer, and J. J. Fallesen. 1989. "Computer-Mediated Group Processes in Distributed Command and Control Systems: Dyad Shared Work." ARI Technical Report 842 (Jun). U.S. Army Research Institute, Ft. Leavenworth, KS.

Losada, M. and S. Markovitch. 1990. "Group Analyzer: A System for Dynamic Analysis of Group Interaction." *Proceedings of the 23rd Hawaii International Conference on System Sciences* (pp. 101–110).

McGrath, J. E., and A. B. Hollingshead. 1994. *Groups Interacting with Technology: Ideas, Evidence, Issues, and an Agenda.* Sage Publications, Thousand Oaks, CA.

Quinn, M. L. 1995. "Design and Evaluation of Tools to Support C2W Team Activities." Pacific Science & Engineering Group, Inc., San Diego, CA.

Quinn, M.L. & Gwynne, J.W. (25 January 1996). "Evaluating Collaborative Technologies for Command and Control Warfare Teams: Support Materials." San Diego: Pacific Science & Engineering Group, Inc., San Diego, CA. Unpublished report available from the author.

Sanderson, P. M. 1994. *MacSHAPA version 1.0.2 Manual.* Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign.

Stubler, W. F., and J. M. O'Hara. 1995. "Group-View Displays: Functional Characteristics and Review Criteria." Prepared for U.S. Nuclear Regulatory Commission, Office of Nuclear Reactor Regulation. Contract No. DE-AC02-76-CH00016 (Jan).

APPENDIX A

SITUATION ASSESSMENT REPORT (SITREP) FORM

Situation Assessment

1. Describe the disposition and capabilities of the threat forces (suspect and hostile). This description may include: track numbers, geographical reference, unit names, installations, etc.

2. Describe the ongoing activities that are most threatening to US forces or to US plans and intentions.

3. What activities do you anticipate monitoring during the next phase? Why might they be important?

4. What changes, if any, do you recommend to the tactical plan?

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1996	Final: October 1994 – January 1996

4. TITLE AND SUBTITLE

EVALUATING COLLABORATIVE TECHNOLOGIES FOR COMMAND AND CONTROL TEAMS

5. FUNDING NUMBERS

N66001–90–D–0127
PE: 0602233N

6. AUTHOR(S)

J. W. Gwynne, R. W. Obermayer, J. M. Linville, N. J. Heacox, Pacific Science and Engineering Group, Inc.
B. A. Feher, R. J. Smillie, Naval Command, Control and Ocean Surveillance Center RDT&E Division;

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Pacific Science and Engineering Group, Inc.
6310 Greenwich Drive
San Diego, CA 92122–5901

Naval Command, Control and Ocean Surveillance Center
RDT&E Division
San Diego, California 92152–5000

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 1728

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution is unlimited.	

13. ABSTRACT *(Maximum 200 words)*

This document describes the U. S. Navy's initial demonstration of the collaborative technology assessment laboratory (COLAB) during a distributed situation assessment task. Five command and control warfare (C^2W) teams, each consisting of three naval officers and a commander, participated in an exercise scenario involving four vignettes that required the team to generate a situation assessment report (SITREP) based on shared information. The COLAB compared the quality, completeness, and consistency of team performance based on collaborative technology with more traditional voice and text communications methods. No advantage could be attributed solely to the collaborative technology. Differences in the frequency and pattern of communications across the traditional and collaborative methods created nonmeaningful influences on the major outcome measures. However, these results led to improved understanding of team collaborative behavior and the creation of a systematic approach for assessing the role of collaborative technology on team performance. As part of this approach, the COBOL developed a set of analytic techniques for evaluating data collected in collaborative studies. These techniques can be applied in future experiments using the COLAB to evaluate collaborative processes and technologies in Navy C^2W operations.

14. SUBJECT TERMS

Mission area: Command and Control
collaborative technology assessment
command and control warfare
timeline analysis
situation assessment reports (SITREPs)

15. NUMBER OF PAGES

52

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b. TELEPHONE *(include Area Code)*	21c. OFFICE SYMBOL
B. Feher	(619) 553-9226	Code 44210

INITIAL DISTRIBUTION

Code D0012	Patent Counsel	(1)
Code D0271	Archive/Stock	(6)
Code D0274	Library	(2)
Code D0271	D. Richter	(1)
Code D404	S. L. Hearold	(1)
Code D404	C. H. Mirabile	(1)
Code D4123	L. T. Duffy	(1)
Code D4204	R. N. Griffin	(1)
Code D44	J. D. Grossman	(1)
Code D441	R. Flemming	(1)
Code D441	G. A. Osga	(1)
Code D4402	R. F. Yturralde	(1)
Code D44209	J. W. Broyles	(1)
Code D44209	J. M. Linville	(5)
Code D44209	R. W. Obermayer	(5)
Code D44210	C. Englund	(1)
Code D44210	S. T. Holte	(1)
Code D44210	J. Morrison	(1)
Code D44210	S. I. Sander	(1)
Code D44210	G. Seymour	(1)
Code D44210	B. Feher	(12)
Code D44210	R. J. Smillie	(5)
Code D834	D. R. Lambert	(1)

Defense Technical Information Center
Fort Belvoir, VA 22060–6218 (4)

NCCOSC Washington Liaison Office
Washington, DC 20363–5100

Center for Naval Analyses
Alexandria, VA 22302–0268

Navy Acquisition, Research and Development
Information Center (NARDIC)
Arlington, VA 22244–5114

GIDEP Operations Center
Corona, CA 91718–8000

Office of Naval Research
Arlington, VA 22217–5660 (6)

Pacific Science & Engineering Group, Inc.
San Diego, CA 92122–5918 (15)

Dr. James A. Ballas
Arlington, VA 22207

Defence Research Agency
Farnborough, HANTS GU14 6TD,
UNITED KINGDOM

University of Arizona
Tucson, AZ 85721–0108

Defence Research Agency Maritime Division
Portsmouth, HANTS P06 4AA,
UNITED KINGDOM (2)

U.S. Naval Air Warfare Center–TSD
Orlando, FL 32768 (2)

Mr. Paul Chatelier
Springfield, VA 22153

San Diego State University
San Diego, CA 92182

Naval Postgraduate School
Monterey, CA 93943–5143 (3)

Dr. Robert S. Kennedy
Winter Park, FL 32792

Naval Underwater Systems Center
Newport, RI 02841

Alphatech, Inc.
Burlington, MA 01803

FU VI Neurologie & Psychiatrie
Keil, GERMANY

George Mason University
Fairfax, VA 22202–1609

CAPT Michael G. Lilienthal
Arlington, VA 22202–1609

Embassy of Australia
Washington, DC 20036–2273

Honeywell Technology Center
Minneapolis, MN 55418

Defence Science & Technology Organization
Canberra, ACT 2600, AUSTRALIA (2)

Temple University
Philadelphia, PA 19122

Defence Research Agency
KENT, TN 14 7BP,
UNITED KINGDOM

Ministry of Defence
London, SW1A 2HB,
UNITED KINGDOM

Dr. Karen Takle Quinn
Los Altos, CA 94013–0277

National Defense University
Washington, DC 20319–6000

University of Minnesota
Minneapolis, MN 55455